



## RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/667,966A  
Source: IFWO  
Date Processed by STIC: 9-28-04

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER  
VERSION 4.2 PROGRAM ACCESSIBLE THROUGH THE U.S. PATENT AND  
TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker/chkrnote.htm>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry, Federal Express, United Parcel Service, or other delivery service (EFFECTIVE 06/05/04):  
U.S. Patent and Trademark Office, 220 20<sup>th</sup> Street S., Customer Window, Mail Stop Sequence, Crystal Plaza Two, Lobby,  
Room 1B03, Arlington, VA 22202

Revised 05/17/04

# Raw Sequence Listing Error Summary

## ERROR DETECTED

## SUGGESTED CORRECTION

SERIAL NUMBER:

10/667,966A

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1 Wrapped Nucleics  
Wrapped Aminos The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2 Invalid Line Length The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3 Misaligned Amino  
Numbering The numbering under each 5<sup>th</sup> amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4 Non-ASCII The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5 Variable Length Sequence(s) contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6 PatentIn 2.0  
"bug" A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s). Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7 Skipped Sequences  
(OLD RULES) Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence:  
(2) INFORMATION FOR SEQ ID NO: X (insert SEQ ID NO where "X" is shown)  
(i) SEQUENCE CHARACTERISTICS (Do not insert any subheadings under this heading)  
(xi) SEQUENCE DESCRIPTION SEQ ID NO: X (insert SEQ ID NO where "X" is shown)  
This sequence is intentionally skipped.  
  
Please also adjust the "(ii) NUMBER OF SEQUENCES" response to include the skipped sequences.
- 8 Skipped Sequences  
(NEW RULES) Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence:  
<210> sequence id number  
<400> sequence id number  
000
- 9 Use of n's or Xaa's  
(NEW RULES) Use of n's and/or Xaa's have been detected in the Sequence Listing.  
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.  
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 10 Invalid <213>  
Response Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence.
- 11 Use of <220>  
Sequence(s) missing the <220>-<223> section and associated numeric identifiers. 3 responses:  
Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.  
(See "Federal Register," 00/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12 PatentIn 2.0  
"bug" Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13 Misuse of n/Xaa "n" can only represent a single nucleotide; "Xaa" can only represent a single amino acid.



IFWO

## RAW SEQUENCE LISTING

DATE: 09/28/2004

PATENT APPLICATION: US/10/667,966A

TIME: 11:36:30

Input Set : A:\US10667966.ST25.txt

Output Set: N:\CRF4\09282004\J667966A.raw

3 <110> APPLICANT: Xie, Dong  
 4 Jiang, He  
 6 <120> TITLE OF INVENTION: Peptide Derivative Fusion Inhibitors of HIV Infection  
 8 <130> FILE REFERENCE: 63024.000002  
 10 <140> CURRENT APPLICATION NUMBER: 10/667,966A  
 11 <141> CURRENT FILING DATE: 2003-09-23  
 13 <150> PRIOR APPLICATION NUMBER: 60/412,797  
 14 <151> PRIOR FILING DATE: 2002-09-24  
 16 <160> NUMBER OF SEQ ID NOS: 15  
 18 <170> SOFTWARE: PatentIn version 3.2  
 20 <210> SEQ ID NO: 1

21 &lt;211&gt; LENGTH: 44

22 &lt;212&gt; TYPE: PRT

23 &lt;213&gt; ORGANISM: Artificial sequence

25 &lt;220&gt; FEATURE:

26 &lt;223&gt; OTHER INFORMATION: FB005 peptide sequence

28 &lt;400&gt; SEQUENCE: 1

30 Ser Leu Glu Gln Ile Trp Asn Asn Met Thr Trp Glu Glu Trp Asp Arg

31 1 .5 10 15

34 Glu Ile Asn Asn Tyr Thr Glu Leu Ile His Glu Leu Ile Glu Glu Ser

35 20 25 30

38 Gln Asn Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu

39 35 40

42 &lt;210&gt; SEQ ID NO: 2

43 &lt;211&gt; LENGTH: 34

44 &lt;212&gt; TYPE: PRT

45 &lt;213&gt; ORGANISM: Artificial sequence

47 &lt;220&gt; FEATURE:

48 &lt;223&gt; OTHER INFORMATION: FB006 peptide sequence

50 &lt;400&gt; SEQUENCE: 2

52 Trp Glu Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Lys Leu Ile His

53 1 5 10 15

56 Glu Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln Glu

57 20 25 30

60 Leu Leu

64 &lt;210&gt; SEQ ID NO: 3

65 &lt;211&gt; LENGTH: 39

66 &lt;212&gt; TYPE: PRT

67 &lt;213&gt; ORGANISM: Artificial sequence

69 &lt;220&gt; FEATURE:

70 &lt;223&gt; OTHER INFORMATION: T-1249 peptide sequence

72 &lt;400&gt; SEQUENCE: 3

74 Trp Gln Glu Trp Glu Gln Lys Ile Thr Ala Leu Leu Glu Gln Ala Gln

Does Not Comply  
 Corrected Diskette Needed

PLS explain source of genetic material.  
 Invalid Response

PLS explain source of genetic material.  
 INVALID Response

PLS see item # 11 on error Summary Sheet.

## RAW SEQUENCE LISTING

DATE: 09/28/2004

PATENT APPLICATION: US/10/667,966A

TIME: 11:36:30

Input Set : A:\US10667966.ST25.txt

Output Set: N:\CRF4\09282004\J667966A.raw

```

75 1           5           10           15
78 Ile Gln Gln Glu Lys Asn Glu Tyr Glu Leu Gln Lys Leu Asp Lys Trp
79           20           25           30
82 Ala Ser Leu Trp Glu Trp Phe
83           35
86 <210> SEQ ID NO: 4
87 <211> LENGTH: 36
88 <212> TYPE: PRT
89 <213> ORGANISM: Artificial sequence SAME error
91 <220> FEATURE:
92 <223> OTHER INFORMATION: T-20 peptide sequence
94 <400> SEQUENCE: 4
96 Tyr Thr Ser Leu Ile His Ser Leu Ile Glu Ser Gln Asn Gln Gln
97 1           5           10           15
100 Glu Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu
101           20           25           30
104 Trp Asn Trp Phe
105           35
108 <210> SEQ ID NO: 5
109 <211> LENGTH: 34
110 <212> TYPE: PRT
111 <213> ORGANISM: Artificial sequence SAME error
113 <220> FEATURE:
114 <223> OTHER INFORMATION: C-34 peptide sequence
116 <400> SEQUENCE: 5
118 Trp Met Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Ser Leu Ile His
119 1           5           10           15
122 Ser Leu Ile Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln Glu
123           20           25           30
126 Leu Leu
130 <210> SEQ ID NO: 6
131 <211> LENGTH: 34
132 <212> TYPE: PRT
133 <213> ORGANISM: Artificial sequence
135 <220> FEATURE:
136 <223> OTHER INFORMATION: SIV C-34 peptide sequence
138 <400> SEQUENCE: 6
140 Trp Gln Glu Trp Glu Arg Lys Val Asp Phe Leu Glu Glu Asn Ile Thr
141 1           5           10           15
144 Ala Leu Leu Glu Glu Ala Gln Ile Gln Gln Glu Lys Asn Met Tyr Glu
145           20           25           30
148 Leu Gln
152 <210> SEQ ID NO: 7
153 <211> LENGTH: 34
154 <212> TYPE: PRT
155 <213> ORGANISM: Artificial sequence SAME error
157 <220> FEATURE:
158 <223> OTHER INFORMATION: FB066 peptide sequence
160 <400> SEQUENCE: 7

```

## RAW SEQUENCE LISTING

DATE: 09/28/2004

PATENT APPLICATION: US/10/667,966A

TIME: 11:36:30

Input Set : A:\US10667966.ST25.txt

Output Set: N:\CRF4\09282004\J667966A.raw

162 Trp Glu Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Lys Leu Ile His  
 163 1 5 10 15  
 166 Glu Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Glu Asn Glu Gln Glu  
 167 20 25 30  
 170 Leu Leu

174 &lt;210&gt; SEQ ID NO: 8

175 &lt;211&gt; LENGTH: 44

176 &lt;212&gt; TYPE: PRT

177 &lt;213&gt; ORGANISM: Artificial sequence

SAME ERROR

179 &lt;220&gt; FEATURE:

180 &lt;223&gt; OTHER INFORMATION: FB005M peptide sequence

183 &lt;220&gt; FEATURE:

184 &lt;221&gt; NAME/KEY: MISC\_FEATURE

185 &lt;222&gt; LOCATION: (23)..(23)

186 &lt;223&gt; OTHER INFORMATION: Xaa represents a Lysine residue derivatized with a maleimide moiety.

189 &lt;400&gt; SEQUENCE: 8

191 Ser Leu Glu Gln Ile Trp Asn Asn Met Thr Trp Glu Glu Trp Asp Arg

192 1 5 10 15

W--&gt; 195 Glu Ile Asn Asn Tyr Thr Xaa Leu Ile His Glu Leu Ile Glu Glu Ser

196 20 25 30

199 Gln Asn Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu

200 35 40

203 &lt;210&gt; SEQ ID NO: 9

204 &lt;211&gt; LENGTH: 45

205 &lt;212&gt; TYPE: PRT

206 &lt;213&gt; ORGANISM: Artificial sequence

SAME ERROR

208 &lt;220&gt; FEATURE:

209 &lt;223&gt; OTHER INFORMATION: FB005CM peptide sequence

212 &lt;220&gt; FEATURE:

213 &lt;221&gt; NAME/KEY: MISC\_FEATURE

214 &lt;222&gt; LOCATION: (45)..(45)

215 &lt;223&gt; OTHER INFORMATION: Xaa represents a Lysine residue derivatized with a maleimide moiety.

218 &lt;400&gt; SEQUENCE: 9

220 Ser Leu Glu Gln Ile Trp Asn Asn Met Thr Trp Glu Glu Trp Asp Arg

221 1 5 10 15

224 Glu Ile Asn Asn Tyr Thr Glu Leu Ile His Glu Leu Ile Glu Glu Ser

225 20 25 30

W--&gt; 228 Gln Asn Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Xaa

229 35 40 45

232 &lt;210&gt; SEQ ID NO: 10

233 &lt;211&gt; LENGTH: 34

234 &lt;212&gt; TYPE: PRT

235 &lt;213&gt; ORGANISM: Artificial sequence

SAME ERROR

237 &lt;220&gt; FEATURE:

238 &lt;223&gt; OTHER INFORMATION: FB006M peptide sequence

241 &lt;220&gt; FEATURE:

242 &lt;221&gt; NAME/KEY: MISC\_FEATURE

## RAW SEQUENCE LISTING

DATE: 09/28/2004

PATENT APPLICATION: US/10/667,966A

TIME: 11:36:30

Input Set : A:\US10667966.ST25.txt

Output Set: N:\CRF4\09282004\J667966A.raw

243 &lt;222&gt; LOCATION: (13)..(13)

244 &lt;223&gt; OTHER INFORMATION: Xaa represents a Lysine residue derivatized with a maleimide moiety.

245

247 &lt;400&gt; SEQUENCE: 10

W--&gt; 249 Trp Glu Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Xaa Leu Ile His

250 1 5 10 15

253 Glu Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Trp Glu

254 20 25 30

257 Leu Leu

261 &lt;210&gt; SEQ ID NO: 11

262 &lt;211&gt; LENGTH: 35

263 &lt;212&gt; TYPE: PRT

264 &lt;213&gt; ORGANISM: Artificial sequence

SAME ERROR

266 &lt;220&gt; FEATURE:

267 &lt;223&gt; OTHER INFORMATION: FB007M peptide sequence

270 &lt;220&gt; FEATURE:

271 &lt;221&gt; NAME/KEY: MISC\_FEATURE

272 &lt;222&gt; LOCATION: (35)..(35)

273 &lt;223&gt; OTHER INFORMATION: Xaa represents a Lysine residue derivatized with a maleimide moiety.

276 &lt;400&gt; SEQUENCE: 11

278 Trp Glu Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Glu Leu Ile His

279 1 5 10 15

282 Glu Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln Glu

283 20 25 30

W--&gt; 286 Leu Leu Xaa

287 35

290 &lt;210&gt; SEQ ID NO: 12

291 &lt;211&gt; LENGTH: 39

292 &lt;212&gt; TYPE: PRT

293 &lt;213&gt; ORGANISM: Artificial sequence

SAME ERROR

295 &lt;220&gt; FEATURE:

296 &lt;223&gt; OTHER INFORMATION: FB010M peptide sequence

299 &lt;220&gt; FEATURE:

300 &lt;221&gt; NAME/KEY: MISC\_FEATURE

301 &lt;222&gt; LOCATION: (13)..(13)

302 &lt;223&gt; OTHER INFORMATION: Xaa represents a Lysine residue derivatized with a maleimide moiety.

305 &lt;400&gt; SEQUENCE: 12

W--&gt; 307 Trp Gln Glu Trp Glu Gln Lys Ile Thr Ala Leu Leu Xaa Gln Ala Gln

308 1 5 10 15

311 Ile Gln Gln Glu Lys Asn Glu Tyr Glu Leu Gln Lys Leu Asp Lys Trp

312 20 25 30

315 Ala Ser Leu Trp Glu Trp Phe

316 35

319 &lt;210&gt; SEQ ID NO: 13

320 &lt;211&gt; LENGTH: 40

321 &lt;212&gt; TYPE: PRT

322 &lt;213&gt; ORGANISM: Artificial sequence

SAME ERROR

## RAW SEQUENCE LISTING

DATE: 09/28/2004

PATENT APPLICATION: US/10/667,966A

TIME: 11:36:30

Input Set : A:\US10667966.ST25.txt

Output Set: N:\CRF4\09282004\J667966A.raw

324 <220> FEATURE:  
 325 <223> OTHER INFORMATION: FB010KM peptide sequence *SAME error*  
 328 <220> FEATURE:  
 329 <221> NAME/KEY: MISC\_FEATURE  
 330 <222> LOCATION: (40)..(40)  
 331 <223> OTHER INFORMATION: Xaa represents a Lysine residue derivatized with a maleimide  
 332 moiety.  
 334 <400> SEQUENCE: 13  
 336 Trp Gln Glu Trp Glu Gln Lys Ile Thr Ala Leu Ile Glu Gln Ala Gln  
 337 1 5 10 15  
 340 Ile Gln Gln Glu Lys Asn Glu Tyr Glu Leu Gln Lys Leu Asp Lys Trp  
 341 20 25 30  
 W--> 344 Ala Ser Leu Trp Glu Trp Phe Xaa  
 345 35 40  
 348 <210> SEQ ID NO: 14  
 349 <211> LENGTH: 34  
 350 <212> TYPE: PRT  
 351 <213> ORGANISM: Artificial sequence *SAME ERROR*  
 353 <220> FEATURE:  
 354 <223> OTHER INFORMATION: FB066M peptide sequence  
 357 <220> FEATURE:  
 358 <221> NAME/KEY: MISC\_FEATURE  
 359 <222> LOCATION: (13)..(13)  
 360 <223> OTHER INFORMATION: Xaa represents a Lysine residue derivatized with a maleimide  
 361 moiety.  
 363 <400> SEQUENCE: 14  
 W--> 365 Trp Glu Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Xaa Leu Ile His  
 366 1 5 10 15  
 369 Glu Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Glu Asn Glu Gln Glu  
 370 20 25 30  
 373 Leu Leu  
 377 <210> SEQ ID NO: 15  
 378 <211> LENGTH: 35  
 379 <212> TYPE: PRT  
 380 <213> ORGANISM: Artificial sequence *SAME error*  
 382 <220> FEATURE:  
 383 <223> OTHER INFORMATION: FB066KM peptide sequence  
 386 <220> FEATURE:  
 387 <221> NAME/KEY: MISC\_FEATURE  
 388 <222> LOCATION: (35)..(35)  
 389 <223> OTHER INFORMATION: Xaa represents a Lysine residue derivatized with a maleimide  
 390 moiety.  
 392 <400> SEQUENCE: 15  
 394 Trp Glu Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Lys Leu Ile His  
 395 1 5 10 15  
 398 Glu Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Glu Asn Glu Gln Glu  
 399 20 25 30  
 W--> 402 Leu Leu Xaa  
 403 35

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/667,966A

DATE: 09/28/2004  
TIME: 11:36:31

Input Set : A:\US10667966.ST25.txt  
Output Set: N:\CRF4\09282004\J667966A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:8; Xaa Pos. 23 ✓

Seq#:9; Xaa Pos. 45 ✓

Seq#:10; Xaa Pos. 13 ✓

Seq#:11; Xaa Pos. 35 ✓

Seq#:12; Xaa Pos. 13 ✓

Seq#:13; Xaa Pos. 40 ✓

Seq#:14; Xaa Pos. 13 ✓

Seq#:15; Xaa Pos. 35 ✓



## VERIFICATION SUMMARY

DATE: 09/28/2004

PATENT APPLICATION: US/10/667,966A

TIME: 11:36:31

Input Set : A:\US10667966.ST25.txt

Output Set: N:\CRF4\09282004\J667966A.raw

L:195 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:16  
L:228 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9 after pos.:32  
L:249 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10 after pos.:0  
L:286 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11 after pos.:32  
L:307 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12 after pos.:0  
L:344 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13 after pos.:32  
L:365 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14 after pos.:0  
L:402 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15 after pos.:32